

Besov regularity of the uniform empirical process

Gane Samb Lo

Ahmadou Bamba SOW

August 27, 2012

LERSTAD, UFR S.A.T, Université Gaston Berger, BP 234, Saint-Louis, SENEGAL.
Phone: (221) 33 961 23 40. Email : ganessamblo@yahoo.com, ahbsow@gmail.com

Abstract

The paths of Brownian motion have been widely studied in the recent years relatively in Besov spaces $B_{p,\infty}^\alpha$. The results are the same as to the Brownian bridge. In fact these regularities properties are established in some sequence spaces $S_{p,\infty}^\alpha$ using an isomorphism between them and $B_{p,\infty}^\alpha$.

In this note, we are concerned with the regularity of the paths of the continuous version of the uniform empirical process in the space $S_{p,\infty}^\alpha$ and in one of his separable sub space $S_{p,\infty}^{\alpha,0}$ for a suitable choice of α and p .

AMS Classification subject: 1991, 60G17, 60G15, 60F15.

Keywords : empirical process, Brownian motion, Besov spaces.

1 Introduction

Let $U_1, U_2, \dots, U_n, \dots$ be a sequence of i.i.d $\mathcal{U}(0,1)$ random variables. For a fix integer $n \geq 1$ we consider the empirical distribution function \tilde{F}_n of the sample U_1, U_2, \dots, U_n defined by

$$\forall 0 \leq s \leq 1, \quad \tilde{F}_n(s) = \frac{1}{n} \sum_{i=1}^n 1_{(-\infty, s]}(U_i)$$

and for $j \geq 0$, $k = 1, \dots, 2^j$ the triangular sequence

$$\tilde{\alpha}_{jk}^n = 2^{j/2} \left[2 \tilde{\alpha}_n\left(\frac{k-1/2}{2^j}\right) - \tilde{\alpha}_n\left(\frac{k-1}{2^j}\right) - \tilde{\alpha}_n\left(\frac{k}{2^j}\right) \right] \quad (1.1)$$

where $\tilde{\alpha}_n$ is the associated empirical process defined by $\tilde{\alpha}_n(s) = \sqrt{n}(\tilde{F}_n(s) - s)$, $0 \leq s \leq 1$.

Our motivation in the study of this sequence is given by previous works on the regularity of the paths of the Brownian motion in Besov spaces $B_{p,\infty}^\alpha$ given by

$$B_{p,\infty}^\alpha = \{f \in L^p([0,1]) : \sup_t \frac{w_p(f,t)}{t^\alpha} < \infty\}$$

where for any $1 \leq p < \infty$,

$$w_p(f,t) = \sup_{|h| \leq t} \left(\int_{I_h} |f(x-h) - f(x)|^p dx \right)^{1/p}; \quad I_h = \{x \in [0,1], x-h \in [0,1]\}.$$

The space $B_{p,\infty}^\alpha$ endowed with the following norm

$$\|f\|_{p,\infty}^\alpha = \sup\{|f_0|, |f_1|, \sup_j 2^{-j(1/2-\alpha-1/p)} \left(\sum_{k=1}^{2^j} |f_{jk}|^p \right)^{1/p}\}$$

where

$$f_0 = f(0), \quad f_1 = f(1) - f(0), \quad f_{jk} = 2^{j/2} \left[2f\left(\frac{k-1/2}{2^j}\right) - f\left(\frac{k-1}{2^j}\right) - f\left(\frac{k}{2^j}\right) \right]$$

is a Banach space.

It is well known thanks to Ciesielski et al [2], that there exists an isomorphism between such spaces and the Banach spaces of sequences $(S_{p,\infty}^\alpha, \|\cdot\|_{p,\infty}^\alpha)$ defined by

$$\{\mu = (\mu_{jk}, j \geq 0, k = 1, \dots, 2^j) / \|\mu\|_{p,\infty}^\alpha < \infty\}$$

where

$$\|\mu\|_{p,\infty}^\alpha = \sup\{|\mu_0|, |\mu_1|, \sup_j 2^{-j(1/2-\alpha-1/p)} \left(\sum_{k=1}^{2^j} |\mu_{jk}|^p \right)^{1/p}\}.$$

Their subsets $B_{p,\infty}^{\alpha,0}$ (respectively $S_{p,\infty}^{\alpha,0}$) of functions $f \in B_{p,\infty}^\alpha$ (resp of sequences $(\mu_{jk}) \in S_{p,\infty}^\alpha$) such that $w_p(f, t) = o(t^\alpha)$ as $t \rightarrow 0$ (resp $2^{-j(1/2-\alpha-1/p)} \left(\sum_{k=1}^{2^j} |\mu_{jk}|^p \right)^{1/p} \rightarrow 0$ as $j \rightarrow \infty$) are separable Banach spaces.

Thanks to this isomorphism, Roynette [4] proved that for $p \geq 2$ and $\alpha < 1/2$, the Brownian path $(W_t, 0 \leq t \leq 1)$ belongs almost surely in $B_{p,\infty}^\alpha$ but not in $B_{p,\infty}^{\alpha,0}$ by establishing

$$\sup_j \left(2^{-j} \sum_{k=1}^{2^j} |g_{jk}|^p \right)^{1/p} < \infty \quad \text{and} \quad \liminf_{j \rightarrow +\infty} \left(2^{-j} \sum_{k=1}^{2^j} |g_{jk}|^p \right)^{1/p} > 0$$

where for all $j \geq 0$ and $k = 1, \dots, 2^j$, $g_{jk} = 2^{j/2} \left[2W\left(\frac{k-1/2}{2^j}\right) - W\left(\frac{k-1}{2^j}\right) - W\left(\frac{k}{2^j}\right) \right]$.

This result can be extended to the Brownian bridge $b_t = W_t - tW_1$, $t \in [0, 1]$ which is closely related to the uniform empirical process. Moreover Komlós et al [3] show that on a suitable probability space $(\mathcal{A}, \mathbf{P})$, there exists a sequence of i.i.d. $\mathcal{U}(0, 1)$ random variables U_1, U_2, \dots , and a sequence of Brownian bridges $\{b_n(t), 0 \leq t \leq 1\}$ such that almost surely

$$\limsup_{n \rightarrow +\infty} \frac{\sqrt{n}}{\log n} \sup_{0 \leq t \leq 1} |\alpha_n(t) - b_n(t)| < \infty.$$

Our aim is to investigate Roynette's result for the Brownian bridge to the continuous version of the empirical process. We successfully get a result for $1 \leq p \leq 2$ and $\alpha = 1/2$.

2 Empirical process and Besov Spaces

In order to face to the lack of smoothness of the classical empirical process, we first recall the following result :

Lemma 2.1 *For every $n \geq 1$, the empirical distribution process $\tilde{\alpha}_n$ admits a continuous version α_n .*

Proof : It is well known one can express the distribution empirical function $\tilde{F}_n(\cdot)$ in terms of the order statistics $U_1^{(n)} \leq U_2^{(n)} \leq \dots \leq U_n^{(n)}$ of the sample U_1, U_2, \dots, U_n as follows

$$\tilde{F}_n(s) = \begin{cases} 0, & U_1^{(n)} > s, \\ \frac{k}{n}, & U_k^{(n)} \leq s < U_{k+1}^{(n)}, \quad k = 1, 2, \dots, n-1, \\ 1, & U_n^{(n)} \leq s. \end{cases}$$

Let us consider the function $F_n(\cdot)$ defined for every $0 \leq s \leq 1$ by

$$F_n(s) = \tilde{F}_n(U_k^{(n)}) + 2 \left(s - \frac{U_k^{(n)} + U_{k+1}^{(n)}}{2} \right) \left(\frac{\tilde{F}_n(U_{k+1}^{(n)}) - \tilde{F}_n(U_k^{(n)})}{U_{k+2}^{(n)} - U_k^{(n)}} \right),$$

if

$$\frac{U_k^{(n)} + U_{k+1}^{(n)}}{2} \leq s \leq \frac{U_{k+1}^{(n)} + U_{k+2}^{(n)}}{2}.$$

It is easy to see that for every $n \geq 1$,

$$\sup_{0 \leq s \leq 1} |F_n(s) - \tilde{F}_n(s)| \leq \frac{1}{n}. \quad (2.1)$$

As a consequence of (2.1), we deduce that F_n is a continuous version of \tilde{F}_n and the process $\alpha_n(s) = \sqrt{n}(F_n(s) - s)$, $0 \leq s \leq 1$ is a continuous version of the associated empirical process $\tilde{\alpha}_n(s) = \sqrt{n}(\tilde{F}_n(s) - s)$, $0 \leq s \leq 1$. ■

We are now in position to formulate our main results

Theorem 2.1 *For every $n \geq 1$, the process α_n satisfy almost surely*

$$\{\alpha_{jk}^n\}_{j,k} \in S_{2,\infty}^{1/2} \quad \text{and} \quad \{\alpha_{jk}^n\}_{j,k} \notin S_{2,\infty}^{1/2,0}.$$

Proof : Let us consider the triangular sequence given by (1.1) (replacing $\tilde{\alpha}_n$ by α_n), we deduce thanks to the distribution empirical process that

$$\forall j \geq 0, \forall k = 1, \dots, 2^j, \quad \alpha_{jk}^n = \frac{2^{j/2}}{\sqrt{n}} \sum_{i=1}^n Z_{jk}(i)$$

where

$$\forall i = 1, \dots, n, \quad Z_{jk}(i) = Z_{jk}(U_i) = 1_{[\frac{k-1}{2^j}, \frac{k-1/2}{2^j}[}(U_i) - 1_{[\frac{k-1/2}{2^j}, \frac{k}{2^j}[}(U_i).$$

Notice that for any $i = 1, \dots, n$, $Z_{jk}(i) \in \{1, 0, -1\}$ respectively with probability $\frac{k}{2^j}, 0, \frac{k}{2^j}$. We deduce that for any $i = 1, \dots, n$, $Z_{jk}(i)$ is centered random variable with variance 2^{-j} .

Let us define $G_{jk} = |\alpha_{jk}^n|^2 = \frac{2^j}{n} H_{jk}$ where $H_{jk} = \left(\sum_{i=1}^n Z_{jk}(i) \right)^2$.

Using the fact that for any fix k and $i \neq h$ the random variables $Z_{jk}(i)$ and $Z_{jk}(h)$ are independent we deduce that

$$\mathbf{E}(H_{jk}) = \mathbf{E}\left(\sum_{i=1}^n Z_{jk}^2(i) + \sum_{i \neq h}^n Z_{jk}(i)Z_{jk}(h)\right) = \frac{n}{2^j}.$$

which implies in particular $\mathbf{E}(G_{jk}) = 1$.

Futhermore for any $j \geq 0$ and $k = 1, \dots, 2^j$, we have

$$\begin{aligned} H_{jk}^2 &= \left(\sum_{i=1}^n Z_{jk}^2(i)\right)^2 + 2 \sum_{l=1}^n Z_{jk}^2(l) \sum_{i \neq h}^n Z_{jk}(i)Z_{jk}(h) + \left(\sum_{i \neq h}^n Z_{jk}(i)Z_{jk}(h)\right)^2 \\ &= \sum_{i=1}^n Z_{jk}^4(i) + 2 \sum_{i < h}^n Z_{jk}^2(i)Z_{jk}^2(h) + 2 \sum_{i \neq h}^n Z_{jk}^2(i)Z_{jk}^2(h) + A_{jk}^{(1)} + A_{jk}^{(2)} \end{aligned}$$

where for every $j \geq 0$ and $k = 1, \dots, 2^j$,

$$A_{jk}^{(1)} = 2 \sum_{l=1}^n \sum_{i \neq h}^n Z_{jk}^2(l)Z_{jk}(i)Z_{jk}(h) \quad \text{and} \quad A_{jk}^{(2)} = 2 \sum_{l \neq m}^n \sum_{i \neq h}^n Z_{jk}(i)Z_{jk}(h)Z_{jk}(l)Z_{jk}(m)$$

It is easy to see that for every $j \geq 0$ and $k = 1, \dots, 2^j$, $A_{jk}^{(1)}$ is a centered random variable and $A_{jk}^{(2)}$ satisfies

$$\begin{aligned} A_{jk}^{(2)} &= 4 \sum_{i < h}^n \left[\sum_{l < m}^n Z_{jk}(i)Z_{jk}(h)Z_{jk}(l)Z_{jk}(m) \right] \\ &= 4 \sum_{i < h}^n \left[\sum_{(l,m)=(i,h)} Z_{jk}(i)Z_{jk}(h)Z_{jk}(l)Z_{jk}(m) \right] + 4 \sum_{i < h}^n \left[\sum_{(l,m) \neq (i,h)} Z_{jk}(i)Z_{jk}(h)Z_{jk}(l)Z_{jk}(m) \right] \end{aligned}$$

The expectation of the last sum vanish thanks to the independence of the random variables. Hence there exists a constant $c > 0$ which may change from line to line such that $\mathbf{E}(A_{jk}^{(2)}) =$

$c \mathbf{E} \sum_{i < h}^n Z_{jk}^2(i) Z_{jk}^2(h)$. We deduce that for every $j \geq 0$ and $k = 1, \dots, 2^j$,

$$\mathbf{E}(H_{jk}^2) = \sum_{i=1}^n \mathbf{E}(Z_{jk}^4(i)) + c \mathbf{E} \sum_{i < h}^n Z_{jk}^2(i) Z_{jk}^2(h) = \frac{n}{2^j} \left(1 + c \frac{n-1}{2^j}\right)$$

which implies in particular for every $j \geq 0$ and $k = 1, \dots, 2^j$,

$$\text{Var}(G_{jk}) = \frac{2^j}{n} \left[1 + \frac{n(c-1) - c}{2^j}\right].$$

Elsewhere we have

$$\begin{aligned} \text{Var}\left(\sum_{k=1}^{2^j} G_{jk}\right) &= \sum_{k=1}^{2^j} \text{Var}(G_{jk}) + 2 \sum_{1=k < k' \leq 2^j} \frac{2^{2j}}{n^2} \text{cov}(H_{jk} H_{jk'}) \\ H_{jk} H_{jk'} &= \left(\sum_{i=1}^n Z_{jk}(i)Z_{jk'}(i) + \sum_{i \neq h}^n Z_{jk}(i)Z_{jk'}(h)\right)^2 \end{aligned}$$

Notice that for $k \neq k'$, the product $Z_{jk}(i)Z_{jk'}(i)$ is null. This implies for every $j \geq 0$ and $k \neq k' \in \{1, \dots, 2^j\}$,

$$\begin{aligned} H_{jk} H_{jk'} &= \left(\sum_{i \neq h}^n Z_{jk}(i) Z_{jk'}(h) \right)^2 = \sum_{i \neq h}^n \sum_{l \neq m}^n Z_{jk}(i) Z_{jk'}(h) Z_{jk}(l) Z_{jk'}(m) \\ &= \sum_{\text{card}\{i, h\} \cap \{l, m\} = 2} Z_{jk}(i) Z_{jk'}(h) Z_{jk}(l) Z_{jk'}(m) \\ &\quad + \sum_{\text{card}\{i, h\} \cap \{l, m\} < 2} Z_{jk}(i) Z_{jk'}(h) Z_{jk}(l) Z_{jk'}(m) \end{aligned} \quad (2.2)$$

So two cases can be investigated :

if $\text{card}\{i, h\} \cap \{l, m\} < 2$, extracting one random variable $Z_{jk}(i)$, the expectation of the last term in (2.2) is null.

if $\text{card}\{i, h\} \cap \{l, m\} = 2$, we have either $(l = i \text{ and } m = h)$ or $(l = h \text{ and } m = i)$. In this last case the product is equal to zero. It remains

$$\begin{aligned} \mathbf{E}(H_{jk} H_{jk'}) &= \sum_{i \neq h}^n \mathbf{E}[Z_{jk}^2(i) Z_{jk'}^2(h)] = \frac{n(n-1)}{2^{2j}} \\ \text{Var}\left(\sum_{k=1}^{2^j} G_{jk}\right) &= \sum_{k=1}^{2^j} \frac{2^j}{n} \left(1 + \frac{3n-4}{2^j}\right) + 2 \sum_{1=k < k' \leq 2^j} \frac{2^{2j}}{n^2} \left(\frac{n(n-1)}{2^{2j}} - \left(\frac{n}{2^j}\right)^2\right) \\ &= 2^{2j} \varepsilon_{nj}, \quad \text{where} \quad \varepsilon_{nj} = \frac{1}{2^j} \left(3 - \frac{3}{n}\right) \end{aligned}$$

Exploiting Bienaymé-Tchébychev inequality, we obtain the following estimate for every $n \in \mathbf{N}$ and

$$\forall j \geq 0, \quad \mathbf{P}\left(\left|2^{-j} \sum_{k=1}^{2^j} |\alpha_{jk}^n|^2 - 1\right| \geq \frac{1}{2}\right) \leq 4 \varepsilon_{nj}$$

Therefore thanks to Borel-Cantelli lemma, we deduce that for any $n \in \mathbf{N}$

$$\frac{1}{2} \leq 2^{-j} \sum_{k=1}^{2^j} |\alpha_{jk}^n|^2 \leq \frac{3}{2} \quad p.s., \quad j \text{ large enough.}$$

Hence

$$\sup_j \left(2^{-j} \sum_{k=1}^{2^j} |\alpha_{jk}^n|^2\right)^{1/2} < \infty \quad \text{and} \quad \liminf_{j \rightarrow +\infty} \left(2^{-j} \sum_{k=1}^{2^j} |\alpha_{jk}^n|^2\right)^{1/2} > 0 \quad a.s.$$

■

Remark:

1. Theses results show that for any n , the sequence (α_{jk}^n) , $j \geq 0$, $k = 1, \dots, 2^j$ belongs in the space $S_{2,\infty}^{1/2}$ and not in $S_{2,\infty}^{1/2,0}$ a.s.
2. The first result of our theorem can be extended to $1 \leq p \leq 2$ since the L^p norm is increasing in p .

References

- [1] B. Boufousssi (1994). *Espaces de Besov: Caractérisations et applications*, Thèse de l'Université Henri Poincaré, Nancy I, France.
- [2] Z. Ciesielski, B. Roynette, G. Kerkycharian (1993). *Quelques espaces fonctionnels associés à des processus gaussiens*, Studia Mathematica, 107, p. 171-204.
- [3] J. Komlós, M. Major, G. Tusnády (1975). *Weak convergence and embedding*. In colloquia Math.Soc. Janos.Boylai.Limit Theorems of Probability Theory, 149-165. Amsterdam, North-Holland.
- [4] B. Roynette (1993). *Mouvement brownien et espaces de Besov*, Stochastics and Stochastics Reports, 43, p. 221-260.